

Case-Based Reasoning Untuk Diagnosa Penyakit Respirologi Anak Menggunakan *Similaritas Simple Mathcing Coefficient*

Tursina

Prodi Teknik Informatika, Fakultas Teknik Universitas Tanjungpura
Jl. Ahmad Yani, Pontianak – 78124
e-mail: tursina15@yahoo.com

Abstract– Penelitian ini mengusulkan salah satu pendekatan dalam diagnosa penyakit Respirogi anak yaitu dengan pendekatan Case-Based Reasoning(CBR). Proses CBR melalui 4 tahap yaitu Retrieve, Reuse, Revise dan Retain. Hal yang terpenting dalam CBR adalah menentukan nilai kemiripan atau similaritas antara kasus-kasus yang tersimpan di basis kasus dengan kasus baru yang akan dicari solusinya. Perhitungan similaritas kasus pada penelitian ini dilakukan dengan cara Simple Matching Coefficient (SMC.)

Kata Kunci : CBR, Similaritas, Simple Matching Coefficien (SMC).

1. Pendahuluan

Case-Based Reasoning (CBR) merupakan penalaran yang bertujuan untuk menyelesaikan permasalahan baru dengan cara mengadapasi solusi-solusi yang terdapat pada kasus-kasus sebelumnya yang mempunyai permasalahan yang mirip dengan kasus yang baru [1]. Sumber pengetahuan utama sistem CBR adalah berdasarkan kasus-kasus yang telah ada atau yang telah tersimpan didalam basis kasus [2].

Kasus-kasus dapat diperoleh dari pengalaman seseorang atau pengalaman seorang pakar dibidangnya. Implementasi CBR sudah banyak dilakukan diberbagai bidang, salah satunya di bidang kedokteran. Dibidang kedokteran CBR banyak digunakan untuk mendukung diagnosa gangguan pada suatu penyakit.

Pada penelitian ini dilakukan diagnosa penyakit yang sering dialami anak-anak yaitu penyakit Respirologi dengan menggunakan metode *case-based reasoning* (CBR) dan *Simple Matching Coefficient Similarity*. Tujuan dari penelitian ini adalah menerapkan metode *Case-Based Reasoning* untuk mendiagnosa penyakit Respirologi anak dengan mencari tingkat kemiripan (similarity) antara basis kasus dengan kasus baru (kasus yang akan didiagnosa)

Sumber data utama yang diperoleh sebagai penunjang dalam membangun CBR pada sistem ini adalah data rekam medis dari Rumah Sakit. Sistem yang dibuat hanya menggunakan fitur berupa gejala-gejala yang tampak pada pasien. *Input* dari sistem yang dibuat berbentuk biner yaitu 1 (ya) atau menyatakan ada gejala dan 0 (tidak) yang menyatakan tidak ada gejala. Sedangkan perhitungan similaritas kasus menggunakan metode *Simple Matching Coefficient (SMC)* yang

merupakan salah satu cara perhitungan similaritas untuk data biner [3]

2. Case Based Reasoning (CBR)

CBR adalah salah satu metode dengan pendekatan berbasis pengetahuan untuk mempelajari dan memecahkan masalah berdasarkan pengalaman pada masa lalu, hal ini berbeda dengan sistem pakar berbasis aturan yang harus mengetahui cara menyelesaikan suatu masalah[2].

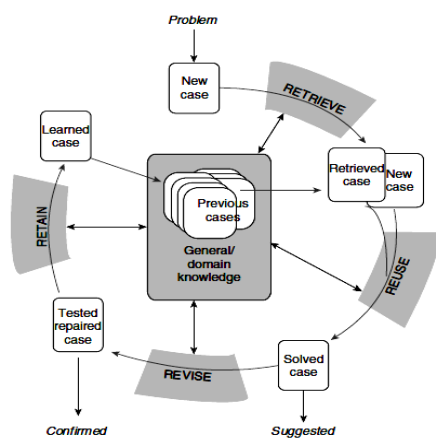
CBR memiliki kemiripan dengan seorang dokter dimana dokter akan meneliti dan mendiagnosa gejala yang terjadi pada pasien, kemudian mencari kasus atau gejala yang pernah dia tangani sebelumnya sehingga dokter memberikan hasil diagnosa dan terapi terhadap pasien seperti halnya yang pernah dilakukan pada pasien terdahulu. [4]

2.1 Metodologi Case-Based Reasoning

CBR harus melakukan beberapa tahapan proses untuk menghasilkan solusi suatu masalah diantaranya adalah mencari tingkat kemiripan kasus baru dengan kasus yang tersimpan. Tahapan proses pada CBR dalam mencari nilai kemiripan dibutuhkan empat (4) tahap, yaitu[1,5] :

1. *Retrieve* (penelusuran) adalah menemukan kembali kasus yang sama atau yang paling mirip dengan kasus baru
2. *Reuse* adalah menggunakan kembali informasi dan pengetahuan dari basis kasus untuk memecahkan masalah kasus baru.
3. *Revise* adalah merevisi atau memperbaiki solusi yang diusulkan.
4. *Retain* adalah menyimpan pengalaman untuk memecahkan masalah yang akan datang kedalam basis kasus.

Tahapan proses CBR dapat dilihat pada Gambar 1.



Gambar 1. Tahapan CBR [1]

Berdasarkan tahapan yang ada dalam CBR, diperlukan tiga langkah utama dalam menentukan solusi, yaitu [7]:

- Membangun basis kasus, yang digunakan sebagai tempat penyimpanan.
- Menentukan fungsi kemiripan (*similarity*), langkah ini digunakan untuk mengenali kesamaan atau kemiripan antara kasus-kasus yang tersimpan dalam basis kasus dengan kasus yang baru.

2.2 Representasi Kasus

CBR tergantung pada struktur dan isi dari koleksi kasus. Suatu kasus dapat diselesaikan dengan memanggil kasus sebelumnya yang sesuai atau cocok dengan kasus baru. Sebuah kasus dapat menjadi sebuah catatan dari kejadian secara khusus terdiri dari [1]:

- a) Permasalahan (*problem*) yang menjelaskan keadaan nyata ketika kasus terjadi.
- b) Solusi, keadaan diperoleh solusi permasalahan.

Pada representasi kasus sistem yang dibangun, setiap kasus dibagi menjadi 3 bagian yaitu:

1. Pasien (yang berisi data-data tentang pasien)
2. Gejala
3. Penyakit.

Pembagian ini dilakukan untuk memudahkan penyimpanan data kasus kedalam basis kasus, serta memudahkan dalam pengambilan data yang sesuai dengan kasus baru. Setiap kasus yang disimpan memiliki tiga bagian yang digunakan dalam memudahkan penyimpanan data kasus.

Dari ketiga bagian tersebut hanya satu bagian yaitu gejala yang hanya digunakan untuk dalam pencarian kasus yang mirip, sedangkan pasien hanya digunakan untuk menyimpan data-data pasien sedangkan penyakit dan terapi merupakan solusi atau *output* dari sistem.

2.3 Penyimpanan

Penyimpanan kasus merupakan aspek penting dalam mendesain CBR. Representasi kasus menggambarkan

gambaran konseptual dalam kasus dan meletakkannya dalam indeks karakteristik kasus. Basis kasus diatur ke dalam sebuah stuktur yang mendukung efisiensi pencarian dan metode penelusuran kembali.

2.4 Penelusuran (*retrieval*) dan Similaritas

Penelusuran yang digunakan pada penelitian ini adalah dengan cara membandingkan fitur antara fitur kasus baru dengan fitur kasus yang ada di basis kasus, kemudian hasil perbandingan tersebut akan dihitung similaritasnya. Fitur yang digunakan adalah berupa gejala-gejala yang tampak pada pasien. Perbandingan fitur menggunakan biner yaitu 1 untuk menyatakan ada gejala dan 0 untuk menyatakan tidak ada gejala.

Perhitungan similaritas digunakan untuk menghasilkan nilai apakah ada kemiripan atau tidak antara kasus baru dengan kasus yang telah ada di basis kasus. Pada penelitian untuk menghitung similaritas antar kasus digunakan metode SMC atau yang dikenal dengan *simple matching similarity* (SMC). SMC adalah satu cara untuk menghitung similaritas dua objek (items) yang bersifat biner [3].

Formula yang digunakan SMC untuk menghitung *similarity* antara dua objek X dan Y adalah sebagai berikut:

Simple matching coefficient (SMC)

$$SMC(X, Y) = \frac{M_{11} + M_{00}}{M_{10} + M_{01} + M_{11} + M_{00}}$$

dimana :

X = Kasus lama

Y = Kasus Baru

M_{11} = Jumlah atribut dimana $X=1$ dan $Y=1$

M_{10} = Jumlah atribut dimana $X=1$ dan $Y=0$

M_{01} = Jumlah atribut dimana $X=0$ dan $Y=1$

M_{00} = Jumlah atribut dimana $X=0$ dan $Y=0$

Kasus baru dikatakan similar (mirip) 100% dengan kasus yang lama apabila nilai similaritas dari $SMC(X, Y)$ sama dengan 1, sedangkan tidak similar apabila nilai $SMC(X, Y)$ sama dengan 0.

Kasus baru (Y) adalah kasus yang akan dicari solusinya dengan cara membandingkan fitur gejala pada setiap kasus lama (X) atau kasus yang tersimpan di basis kasus. Banyaknya biner tiap kasus sesuai dengan banyaknya jumlah gejala keseluruhan yang ada pada basis kasus.

2.5 Adaptasi

Ketika sebuah kasus yang mirip ditemukan, CBR akan mencoba menggunakan kembali solusi yang disarankan dari kasus yang mirip yang telah ditemukan. CBR akan mengadaptasi solusi yang tersimpan dari kasus yang ditemukan kedalam kasus yang ada. Secara tidak langsung CBR sudah melakukan penambahan pengetahuan untuk digunakan di masa yang akan

datang. Beberapa teknik yang digunakan dalam CBR untuk adaptasi antara lain [5]:

1. Tidak ada adaptasi (*Null Adaptation*).
2. Penyesuaian atau pengaturan parameter.
3. *Reinstantiation*.
4. *Derivational replay*.
5. Perbaikam model terpadu.

Pada sistem ini menggunakan *null adaptation* yaitu sistem hanya menerima apapun solusi yang diberikan dari hasil penelusuran kasus yang mirip dengan kasus yang baru.

3. Analisis dan Perancangan Sistem

3.1 Deskripsi Sistem

Sistem yang akan dibangun adalah *Case-Based Reasoning* kasus untuk diagnosa penyakit Respirologi anak. *Input* dari sistem adalah gejala-gejala yang tampak pada anak, sedangkan *output* sistem adalah nama penyakit berdasarkan dari gejala yang di-*input*-kan.

3.2 Akuisisi pengetahuan

Akuisisi pengetahuan adalah proses untuk mengumpulkan data-data pengetahuan dari sumber pengetahuan. Sumber pengetahuan yang akan dijadikan sebagai bahan acuan dalam penelitian ini:

1. Seorang Pakar: Dokter spesialis penyakit anak.
2. *Medical Record* yang berhubungan dengan penyakit paru dan pernafasan anak
3. Buku penunjang penelitian yang berhubungan dengan penyakit Respirologi anak.

3.3 Repesentasi Kasus

Menentukan fitur-fitur dari kasus yang akan dijadikan sebagai basis kasus. Pada penelitian ini hanya mendiagnosa 9 penyakit yang berhubungan dengan Respirologi yaitu:

- Rhinitis
- Influenza
- Bronchitis
- Asma: ringan, sedang, berat.
- Tuberkolosis paru
- Otitis media akut
- Faringitis
- Sindrom croup: ringan, sedang, berat.
- Salesma

3.4 Retrieval dan similaritas

Retrieval yang digunakan dalam penelitian ini adalah membandingkan setiap gejala kasus baru dengan gejala-gejala yang ada pada setiap kasus yang ada di basis kasus, perbandingan tersebut dihitung dengan menggunakan similaritas. Jika nilai basis kasus yang dibandingkan sama atau hampir sama dengan nilai kasus baru maka solusi dari basis kasus tersebut akan disarankan untuk menjadi solusi dari kasus baru.

Nilai similaritas kasus antara 0 sampai dengan 1. Kasus baru dikatakan mirip 100% apabila similaritas antara kasus baru dengan kasus yang ada dalam basis

kasus bernilai 1, sebaliknya jika tidak sama sekali kemiripan akan bernilai 0.

Setiap fitur (gejala) tidak diberikan pembobotan, sedangkan data yang di-*input*-kan pada sistem berbentuk biner 1 atau 0. *Input* 1 (ya) untuk menyatakan bahwa ada gejala dan 0 (tidak) menunjukkan tidak ada gejala. Perhitungan similaritas menggunakan *Simple Matching Coefficient* (SMC)

3.5 Update Kasus

Update kasus akan dilakukan apabila kasus baru yang didiagnosa mempunyai nilai dibawah *threshold* atau tidak berhasil didiagnosa (nilai similaritas 0). Pada sistem ini ditentukan *threshold* sebesar 0.95 yang digunakan sebagai indikator apakah kasus yang baru tersebut akan di-*retain* ke basis kasus atau tidak.

Kasus baru yang di-*update* akan tersimpan di basis kasus, secara langsung sistem akan mendapatkan pengetahuan baru dari kasus baru yang telah di-*update* ke dalam basis kasus.

4. Hasil dan Pembahasan

4.1 Proses Pengisian Basis Kasus

Tahap awal dari penggunaan system CBR adalah pengisian basis kasus. Data-data yang akan dijadikan basis kasus adalah yang berhubungan dengan penyakit paru dan pernafasan pada anak-anak yang berusia antara 3 sampai dengan 10 tahun.

4.2 Proses Diagnosa dan Update Kasus

Diagnosa penyakit dilakukan dengan cara memasukkan data pasien dan gejala-gejala kasus yang akan didiagnosa ke dalam *form* pengisian basis kasus. Ketika gejala pada kasus baru dimasukkan, maka sistem akan mencari tingkat kemiripan dengan kasus-kasus yang tersimpan pada basis kasus. Sistem akan menampilkan hasil diagnosa kasus berupa nilai similaritas.

Solusi kasus baru dinyatakan mirip atau solusi dari kasus yang lama dapat digunakan untuk menyelesaikan kasus yang baru apabila nilai kemiripannya lebih besar atau sama dengan 0.95 (nilai *threshold*).

Kasus yang tidak berhasil didiagnosa maka akan dilakukan *update* kasus. Ada dua kondisi untuk dilakukannya *update* kasus:

- Apabila kasus yang didiagnosa tidak mempunyai kemiripan sama sekali dengan kasus-kasus yang ada di basis kasus (similaritas = 0).
- Apabila kasus yang diagnosa mempunyai nilai kemiripan dibawah 0.95 (dibawah nilai *threshold*), maka akan dilakukan *update* kasus karena derajat kepercayaan dinilai rendah atau tidak terlalu besar.

4.3 Pengujian Sistem

Kasus-kasus yang digunakan dalam pengujian sistem ini adalah kasus-kasus yang diperoleh dari rekam medis sebanyak 250 kasus. Kasus-kasus yang diperoleh tidak semua dimasukkan kedalam basis kasus tetapi terbagi dua bagian yaitu 150 kasus digunakan sebagai basis kasus (yang tersimpan didalam sistem) sedangkan 100 kasus digunakan untuk pengujian sistem.

Uji coba sistem dilakukan dengan 2 cara, yaitu:

1. Pengujian tanpa *threshold*.
Pengujian dilakukan dengan cara mendiagnosa 100 kasus baru terhadap 150 kasus yang ada di basis kasus.
2. Pengujian dengan *threshold*.
 - Pengujian dilakukan terhadap 100 kasus baru terhadap 150 kasus yang terdapat pada basis kasus.
 - Pengujian menggunakan 3 nilai *threshold* yang berbeda-beda yaitu: 0.50, 0.75 dan 0.95.

Pengujian 1 dan 2 akan mendiagnosa kasus baru dengan cara menghitung similaritas antara kasus baru (100 kasus) dengan kasus yang tersimpan di basis kasus (150 kasus).

4.4 Pengujian tanpa Threshold

Pada pengujian ini dilakukan pengujian terhadap 100 kasus baru yang berbeda-beda. Hasil yang diharapkan pada pengujian ini adalah apakah diagnosa yang dilakukan sistem menggunakan metode SMC sudah sesuai dengan hasil dari pakar.

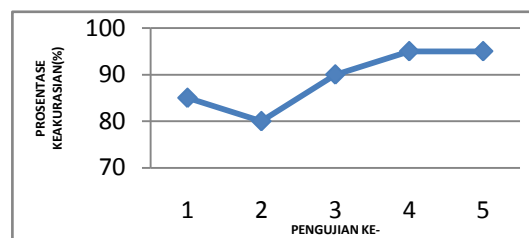
Pengujian dilakukan sebanyak 5 kali. Setiap pengujian dilakukan dengan mendiagnosa 20 kasus baru yang berbeda-beda di setiap pengujian jadi total kasus baru yang akan di uji sebanyak 100 kasus baru. Hasil pengujian 1 terhadap 20 kasus baru dapat dilihat pada Tabel 1. Pada Tabel 2 diperlihatkan nilai similaritas dan nama penyakit berdasarkan perhitungan dengan metode SMC 20 kasus baru yang di-*input*-kan.

Tabel 1. Hasil Diagnosa SMC terhadap Pakar untuk 20 kasus baru yang diuji.

Kasus baru	Similaritas	Nama Penyakit	PAKAR
1	0.97	Salesma	Salesma
2	0.94	Faringitis	Faringitis
3	0.94	Influenza	Influenza
4	0.88	Salesma	Influenza
5	0.97	Asma Ringan	Asma Ringan
6	0.94	TB	Tuberkulosis (TB)
7	1	Sindrom Croup Sedang	Sindrom Croup Sedang
8	0.91	Sindrom Croup Sedang	Sindrom Croup Ringan
---	---	---	---
20	0.94	Faringitis	Faringitis

Dari hasil pengujian 100 kasus baru diperoleh rata-rata prosentase keakurasian diagnosa 100 kasus baru yang diuji menggunakan metode SMC terhadap diagnosa pakar sebesar 89%.

Prosentase keakurasian dapat dilihat pada Gambar 2.



Gambar 2. Prosentase 5 kali pengujian terhadap 100 kasus baru dengan Metode Simple Matching Coefficient (SMC).

Pengujian juga dilakukan dengan menghitung rata-rata berdasarkan banyaknya kasus berdasarkan nama penyakit. Pengujian dilakukan terhadap kasus baru yang ke 20 diperoleh nilai rata-rata similaritas dapat dilihat pada Tabel 2.

Tabel 2. Nilai rata-rata Similaritas berdasarkan penyakit

No	Nama Penyakit	Rata-Rata Similaritas
1	Salesma	0.0642
2	Faringitis	0.8666
3	Influenza	0.7095
4	Sindrom Croup Ringan	0.7857
7	Otitis media Akut	0.8142
..
12	Asma Berat	0.7171

Dari Tabel 2 diperoleh bahwa untuk kasus baru (kasus ke 20) nilai similaritas yang tertinggi kedua metode adalah penyakit Faringitis.

4.5 Pengujian dengan Threshold

Pada pengujian ini dilakukan terhadap 100 kasus baru yang sama seperti pada pengujian sebelumnya (tanpa *threshold*). Hasil yang diperoleh banyaknya kasus yang ada di basis kasus (150 kasus) yang mirip dengan 100 kasus baru yang di-*input*-kan dengan menetapkan *threshold* (0.50, 0.75 dan 0.95). Hasil pengujian dapat dilihat pada Tabel 3.

Tabel 3. Banyaknya kasus yang mirip berdasarkan *threshold*

Kasus Baru	Banyaknya kasus berdasarkan threshold		
	0.5	0.75	0.95
1	150	19	1
2	150	111	0
3	150	107	0
4	150	26	0
..
100	150	19	4

Pada Tabel 3 menunjukkan banyaknya kasus yang berada pada *threshold-threshold* yang telah ditentukan, misalnya pada pengujian untuk kasus 1 terhadap 150 kasus yang di basis kasus diperoleh:

- ✓ Banyaknya kasus yang berada di *threshold* 0.50 sebanyak 150 kasus (100%). Hal ini berarti semua nilai similaritas kasus 1 terhadap 150 kasus yang ada di basis kasus berada di atas nilai *threshold* 0.5.
- ✓ *Threshold* 0.75 sebanyak 19 kasus (12.67% dari 150 kasus)
- ✓ *Threshold* 0.95 sebanyak 1 kasus (0.67% dari 150 kasus)

Dari hasil pengujian dengan *threshold* 0.95 diperoleh kasus-kasus yang berada diatas *threshold* dengan sebanyak 46 kasus dari 100 kasus baru yang diuji.

Prosentase banyaknya kasus berdasarkan *threshold* (0.50, 0.75 dan 0.95) dari 100 kasus baru yang diuji terhadap 150 kasus di basis kasus dapat dilihat pada Tabel 4.

Tabel 4. Prosentase banyaknya kasus berdasarkan *threshold*

No. Kasus	Prosentase Banyaknya kasus (%)		
	SMC		
	0.5	0.75	0.95
1	100	12.67	0.667
2	100	74	0
3	100	71.33	0
4	100	17.33	0
5	100	73.33	1.333
6	100	76	0
7	100	71.33	4
8	100	58	0
9	100	90	0.667
...
100	100	12.67	0.667

5. Kesimpulan

- Gejala-gejala penyakit pada sistem ini digunakan sebagai *problem case* sedangkan nama penyakit digunakan sebagai *solution space*.
- Nilai kemiripan antara kasus baru dengan kasus yang ada di basis kasus bernilai antara 0 atau 1. Nilai 0 berarti menunjukkan kemiripan kasus tidak ada sedangkan kemiripan dikatakan sama apabila kemiripan kasus baru dengan kasus yang ada di basis kasus bernilai 1.
- Proses *retrieval* dilakukan dengan cara membandingkan setiap fitur (gejala) antara kasus baru dengan kasus yang ada di basis kasus, sedangkan perhitungan similaritas dilakukan dengan

metode SMC. Sistem akan menampilkan satu kasus yang mempunyai nilai similaritas tertinggi.

- Pengujian dengan metode SMC terhadap 100 kasus baru mempunyai rata-rata nilai keakurasian sebesar 89%.
- Ada dua kondisi *update* kasus: pertama, kasus baru yang didiagnosa tidak mempunyai kemiripan sama sekali dengan kasus yang terdapat di basis kasus (similaritas bernilai 0). Kedua, kasus baru memiliki kemiripan dengan kasus yang terdapat di basis kasus tetapi mempunyai nilai dibawah 0.95, sehingga derajat kepercayaan hasil diagnosa tidak terlalu besar.

Referensi

- [1] Althoff, K. D. (2001). *Case-Based Reasoning*, Handbook of Software Engineering & Knowledge Engineering (ed. S.K. Chang) Vol 1, World Scientific, Singapore.
- [2] Watson, I. (1997). *Applying Case-Based Reasoning, Technique for Enterprise Systems*, Morgan Kaufmann Publishers.
- [3] Tan, P.N., M. Steinbach and V.Kumar. (2005). *Introduction to Data Mining*, Addison Wesley.
- [4] M. Salem dan Abdel-Badeeh. (2007). *Case Based Reasoning Technology for Medical Diagnosis*, World Academy of Science, Engineering and Technology, <http://www.waset.org/journals/waset/v31/v31-2.pdf>, diakses pada tanggal 17 maret 2010.
- [5] Sankar K. Pal dan Simon. (2004). *Foundations Of Soft Case-Based Reasoning*, Wiley-Interscience.
- [6] Swaboda, W., Zwiebel, F.M., Spitz, R., and Gierl, L. (1994). *A Case based consultation system for postoperative management of liver-transplanted patient, Proceeding of the 12th MIE Lisbon*, IOS Press, Amsterdam, pp.191-195.
- [7] Kusri and Hartati, S. (2011) *Penggunaan Penalaran berbasis Kasus Untuk membangun Basis Pengetahuan dalam Sistem Diagnosa Penyakit*, [Http://dosen.amikom.ac.id/downloads/artikel/full_paper_sr_iti_kusri.pdf](http://dosen.amikom.ac.id/downloads/artikel/full_paper_sr_iti_kusri.pdf), diakses pada tanggal 17 maret 2011.

Biography

Tursina, lahir di Semparuk, Kalimantan Barat, Indonesia pada tanggal 15 Januari 1978. Memperoleh gelar Sarjana Informatika dari Universitas Islam Indonesia (UII), Jogjakarta, Indonesia pada tahun 2001 dan gelar *Master of Computer Science* dari Universitas Gadjah Mada (UGM) Jogjakarta pada tahun 2011, Indonesia. Bidang penelitian saat ini *Case-Based Reasoning*.

